



Linked Open Data

# Semantiek, URI's, LD Applicaties, Tijd en locatie

Workshop, discussie, huidige stand van zaken  
17 april 2014

## Fysiek object

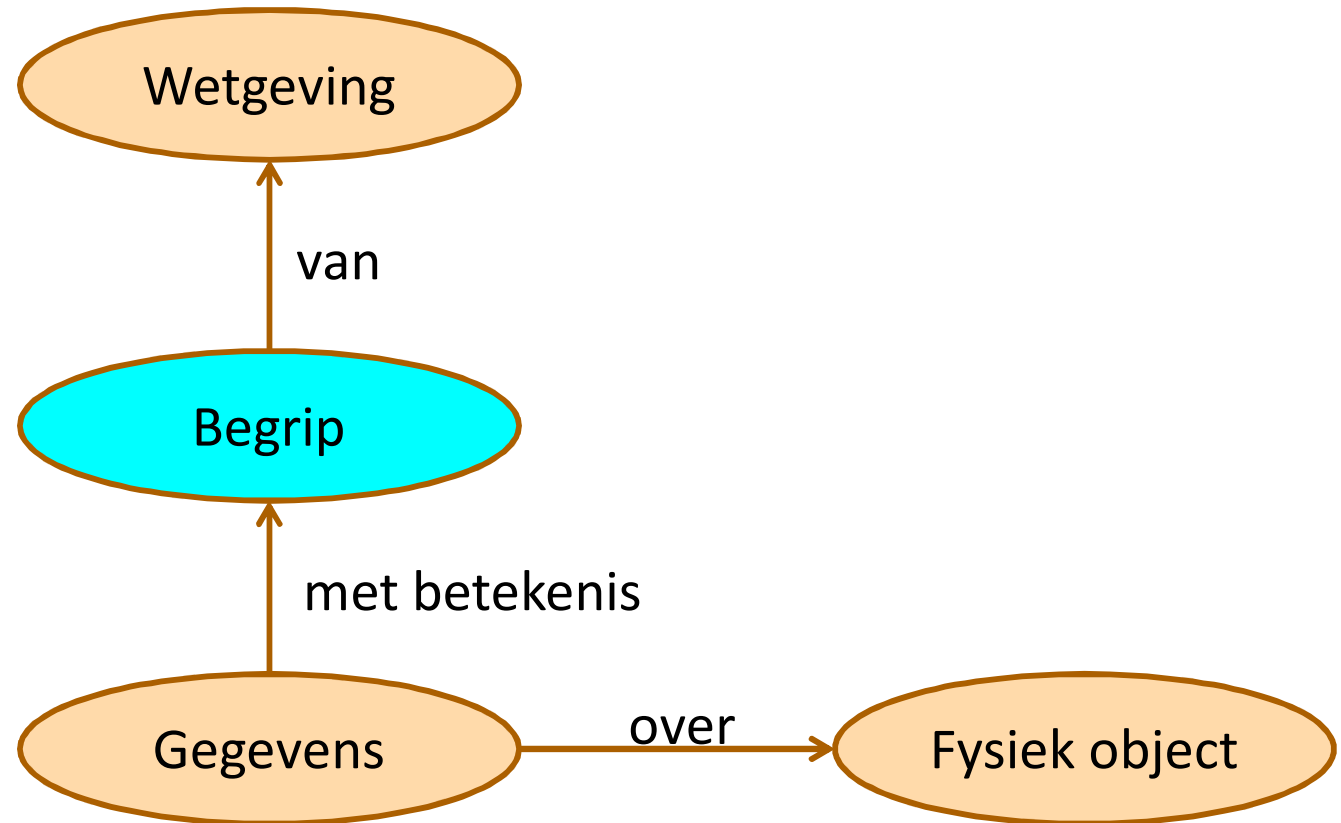


## Wetgeving

**Artikel 1 lid 1 Wet BAG**  
nummeraanduiding: door het bevoegde gemeentelijke orgaan als zodanig toegekende aanduiding van een verblijfsobject, een standplaats of een ligplaats;

## Gegevens

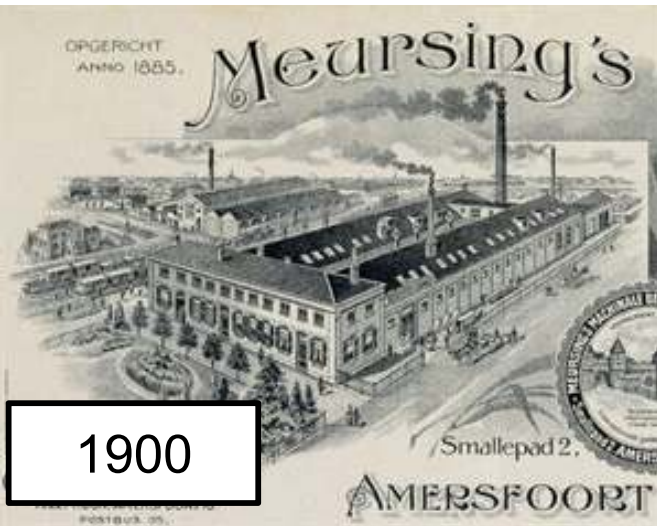
```
<bag:Nummeraanduiding rdf:about="http://bag.kadaster.nl/id/nummeraanduiding/0307200000541769">
  <bag:identificatie>0307200000541769</bag:identificatie>
  <bag:huisnummer>5</bag:huisnummer>
  <bag:officieel>Nee</bag:officieel>
  <bag:postcode>3811 MG</bag:postcode>
  <bag:begindatumTijdvakGeldigheid>13-04-2012</bag:begindatumTijdvakGeldigheid>
  <bag:inOnderzoek>Nee</bag:inOnderzoek>
  <bag:gerelateerdeOpenbareRuimte rdf:resource="http://bag.kadaster.nl/id/openbareruimte/0307300000306884"/>
  <bag:typeAdresseerbaarObject>Verblijfsobject</bag:typeAdresseerbaarObject>
  <bag:bron rdf:resource="http://bag.kadaster.nl/id/document/D4080390/13-04-2012"/>
  <bag:status>Naamgeving uitgegeven</bag:status>
  <rdfs:label>Smallepad 5 3811 MB Amersfoort</rdfs:label>
</bag:Nummeraanduiding>
```

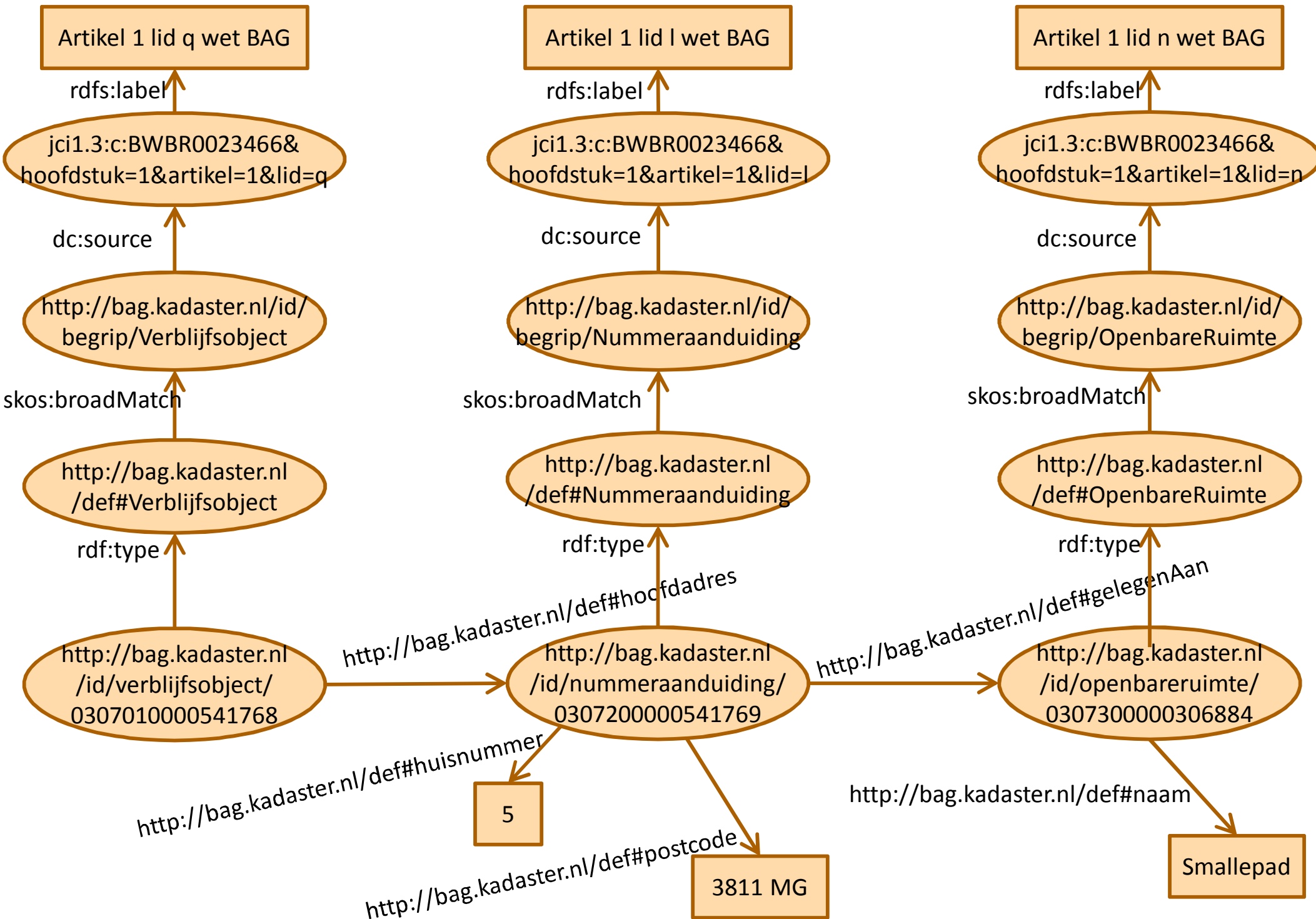


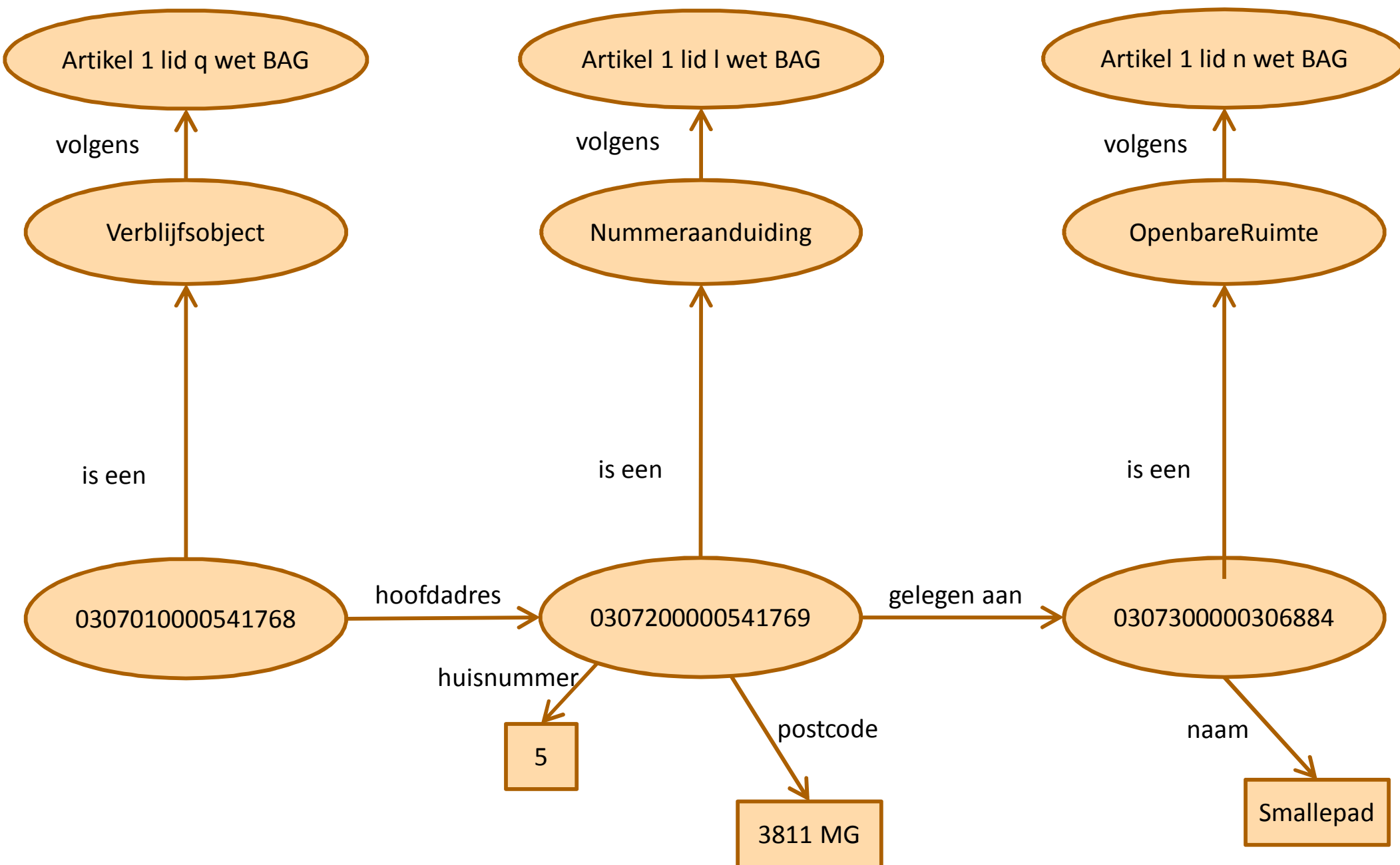
**Linked data als verbindende factor**



Zowel materiele als formele historie:  
de wereld verandert (materieel), onze vastlegging hiervan verandert (formeel)





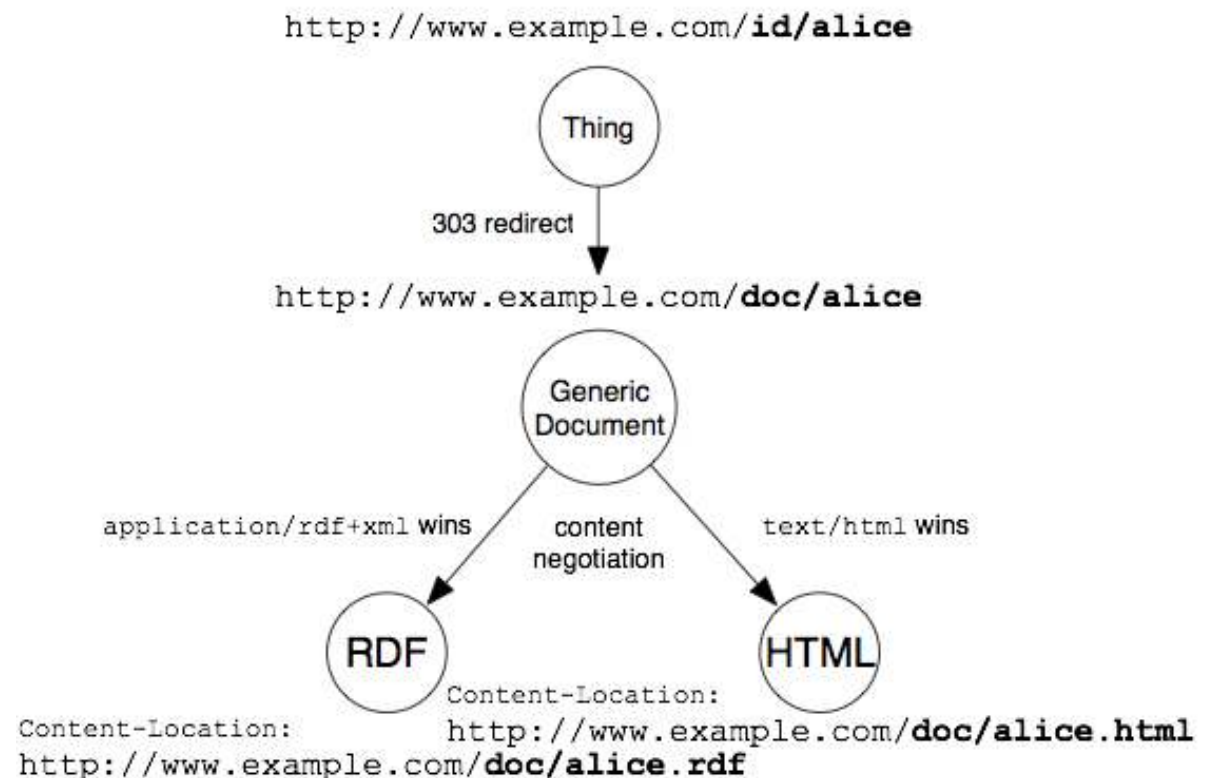


1. Vastlegging van geo- en tijdsaspecten: gegevens *over* of context *van*?
2. Uitvragen van context: onderdeel van de URI strategie?
  - Gebruik van een URI voor de *identificatie* van een resource, en...
  - Gebruik van een URL voor een http request (REST API).
3. URI strategie: URL's of URN's?
  - “Tim zegt”: http-URI's
  - URN's bestaat (en niet de minste): Juriconnect, ECLI, ELI
4. Opslaan van context: in de URI of in de graph?
  - Bijvoorbeeld: bronvermelding, versie, geldigheid, historie, etc, etc.
  - In de URI betekent: URI strategie nodig voor dergelijke aspecten;
  - In de graph betekent: quad-store implementatie



`http://{domain}/{type}/{concept}/{reference}`

- {domain} = {internet domain}/{path}
- {type} = 'id' , 'doc', 'def'
- {concept} = identificatie (voor de mens)
- {reference} = unieke identificatie (voor register / computer)



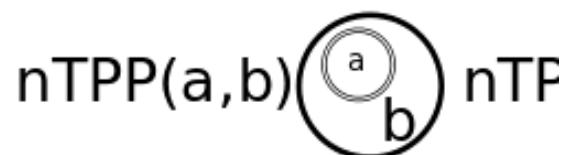
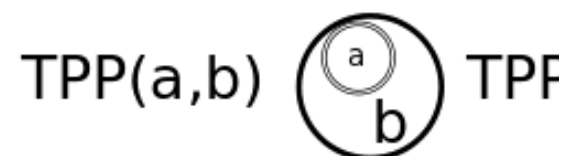
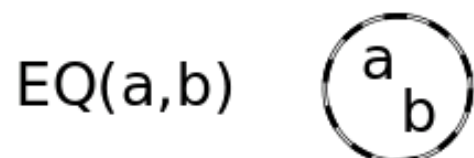
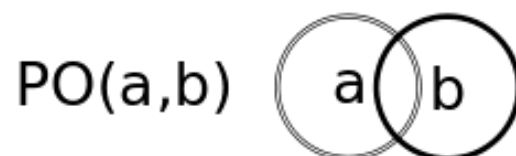
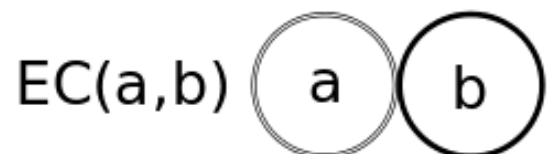
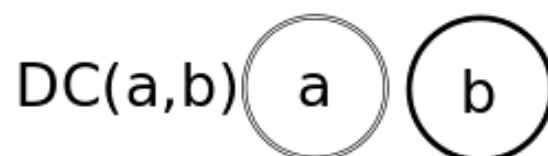


- Geen informatie in de URI!! Ik ben nog steeds erg voor: `http://<uuid>`
- Daarnaast weet ik niet of het slecht is om datum / tijd als versies op te nemen in URI's  
bijv:  
<http://data.overheid.nl/bag/pand/363100/20140406T133400>  
of UUID  
<http://data.overheid.nl/bag/pand/363100/3a615380-bd7f-11e3-a593-0002>
- Ik ben voorstander om zo dicht mogelijk te blijven bij de identifiers zoals ze in de oorspronkelijke registratie bedacht zijn als je URI's gaat bedenken.
- In het meest dramatische geval UUID-iseren (dit werkwoord is nog niet rijp voor de Van Dale) we het gehele RDF naar onleesbare triples:  
<http://data.overheid.nl/bce32533-87ce-436e-a33d-42a4bf50e804>  
liever: <http://data.overheid.nl/bag/pand/363100/versie-20140407-20101223>
- Ik vind het zinvol om de even terug te grijpen op de URI-strategie  
1)Persistentie - 2)Schaalbaarheid - 3)Begrijpelijkheid - 4)Vertrouwen - 5)Machine-leesbaarheid - 6)Menselijke leesbaarheid.  
Het omgaan met tijd en URI's lijkt vooral een afweging tussen 1-2-5 en 3-4-6 oftewel iets wat de computer goed begrijpt of de mens ook nog begrijpt/vertrouwt.

- Het is een aantoonbaar slecht idee om wat voor informatie dan ook in een URI te stoppen.
- Om de herkenbaarheid van de URI te vergroten zou ik de graphs ook een UUID geven.  
`http://{hostname}/{graph-uuid}/{resource-uuid}`  
bag.kadaster.nl/kopen-1234/huis-123456 en (life events)  
bag.kadaster.nl/verkopen-1248/huis-123456
- Mijn persoonlijke inzicht van vandaag is dat een Canonical URI niet bestaat.  
Het is meer van:
  - Als ik de Temporal aspecten in de URI wil opnemen (ipv. op het niveau van de ontology) dan los ik dat op manier A op.
  - Als ik de Provenance aspecten in de URI wil opnemen, dan los ik dat op manier B op.  
Waarbij de ene gegevensset manier A kan gebruiken en de andere gegevensset manier B.
- Hmm, URI strategie, versionering en lifecycles, da's een pittig thema bij elkaar...
- keeping state and identity separated is indeed a good idea.

<http://www.opengeospatial.org/standards/geosparql>

Relation Name	Relation URI
equals	geo:sfEquals
disjoint	geo:sfDisjoint
intersects	geo:sfIntersects
touches	geo:sfTouches
within	geo:sfWithin
contains	geo:sfContains
overlaps	geo:sfOverlaps
crosses	geo:sfCrosses



**SQL:2011 or ISO/IEC 9075:2011**

SQL Language enhancements for temporal data definition and manipulation

- Valid time is the time period during which a fact is true with respect to the real world.
- Transaction time is the time period during which a fact stored in the database is considered to be true.
- Bitemporal data combines both Valid and Transaction Time.

<b>Interval functions</b>	Interval
<b>Timestamptz functions</b>	first - last - prior - next (period p)
<b>Boolean functions</b>	contains - adjacent - overlaps - overleft - overright - isempty equals - before - after
<b>Period functions</b>	period - empty_period - intersect - union - minus

<http://www.strabon.di.uoa.gr/stSPARQL#temporals>

- Spatial Extension Functions
- Topological Relationships
- Topological Relationships (based on minimum bounding boxes)
- Spatial Analysis Functions
- Temporal Extension Functions
  - xsd:boolean strdf:after(strdf:period i1, strdf:period i2): returns true if i1 happens after i2.
  - xsd:boolean strdf:starts(strdf:period i1, strdf:period i2): returns true if i1 starts at the same time as period i2.

Compliant with OGC , SQL11 , GeoSPARQL (almost)

Unfortunately, not a standard (yet).



## Discreet:

[http://wetten.overheid.nl/BWBR0032789/geldigheidsdatum\\_16-04-2014#Artikel5](http://wetten.overheid.nl/BWBR0032789/geldigheidsdatum_16-04-2014#Artikel5)

## Continu:

<http://..weer.nl/id/neerslag/..?>

op plaats (x,y\_ en dag (t)

<http://..brk.nl/id/perceel/6345789/..?>

eigendomsrecht vorig jaar

<http://..bag.nl/id/openbareruimte/4443323/...?> Versie voor correctie dit jaar.

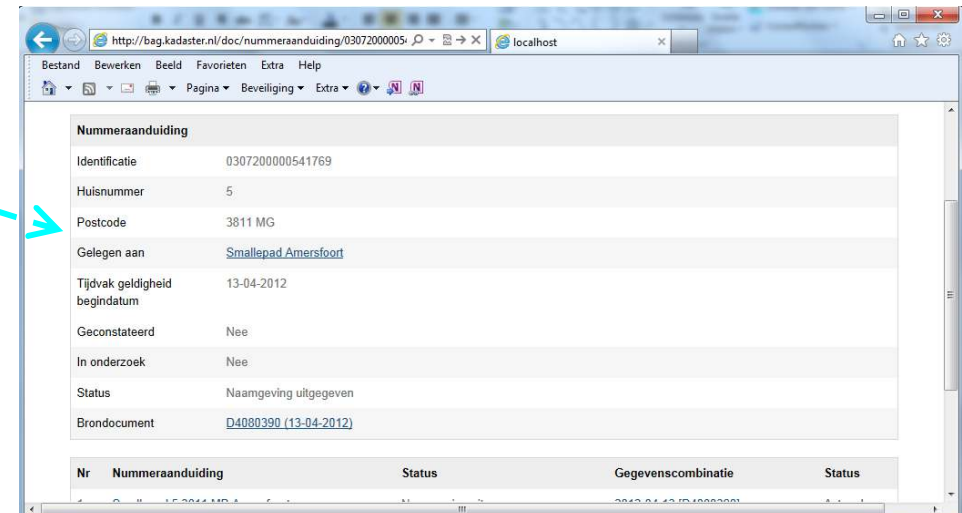
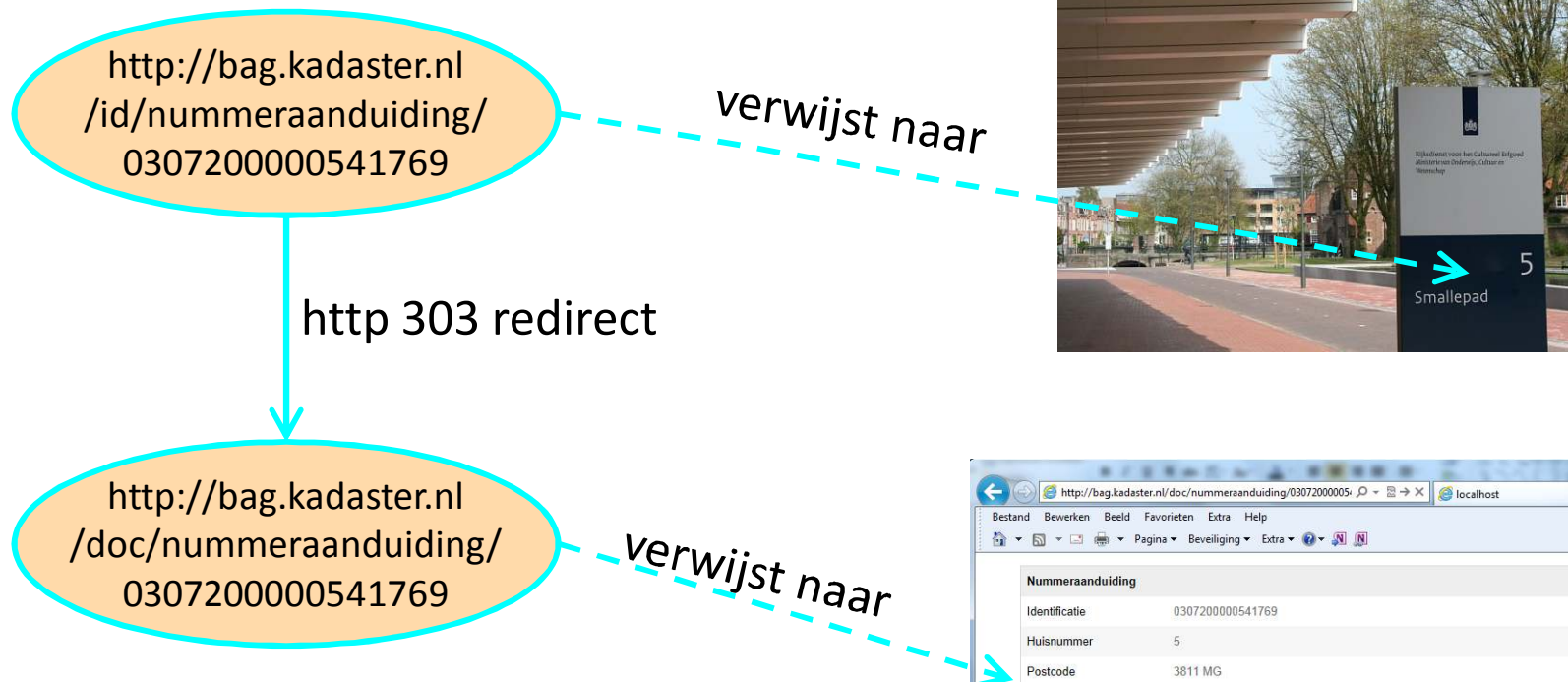
De URI's uit de URI-strategie gaan over het munten van URI's voor “dingen” (abstract of concreet)

<http://bag.kadaster.nl/id/nummeraanduiding/0307200000541769>

verwijst naar



Als zo'n URI een URL is, dan kun je via een http 303 redirect terecht komen op een *pagina* met informatie over dit “ding”



**!** NB: De informatie die je krijgt is afkomstig van dezelfde “eigenaar” als degene die de URI heeft gemunt (in dit geval: bag.kadaster.nl)

“Normale” URL’s zijn geen identificaties van “dingen”, maar manieren om informatie over “iets” te krijgen vanuit een bepaalde context:

- De domeinnaam geeft dan weer *van wie* je de informatie wilt ontvangen
- De rest van de URL geeft eventueel aanwijzingen voor de context.
- Hierover zijn ook afspraken te maken, maar vallen eigenlijk **buiten** een URI-strategie die het doel heeft om URI’s te munten voor identificaties!

Bijvoorbeeld:

<http://mijndomein.nl/about?resource=http%3A%2F%2Fbag.kadaster.nl%2Fid%2Fopenbareruimte%2F0307300000306884&tijdstip=20130101&viewpoint=provincie&bron=gba>

(betekenis: geef informatie aanwezig op **mijndomein.nl** over het “ding” met de URI <http://bag.kadaster.nl/openbareruimte/0307300000306884>, zoals daar bekend is hoe het **gba** hierover denkt, zoals de **provincie** dit graag ziet).

## **“Tim zegt”: gebruik http-URI's**

### *Voordeel van http-URI's (=URL's):*

- Er is geen resolver nodig: je kunt direct informatie ophalen via de URL;
- De URL geeft informatie over het eigenaarschap.

### *Nadeel van http-URI's (=URL's):*

- Een URL vereist ook een server waar informatie is te vinden;
- De eigenaar van een domein kan veranderen: wat betekent dit voor de URL?

### *Voordeel van URN's:*

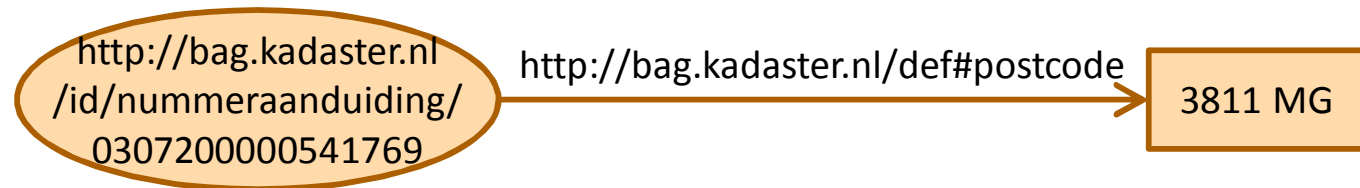
- Te munten zonder dat er ook een server moet zijn;
- Niet afhankelijk van een implementatie.

### *Nadeel van URN's:*

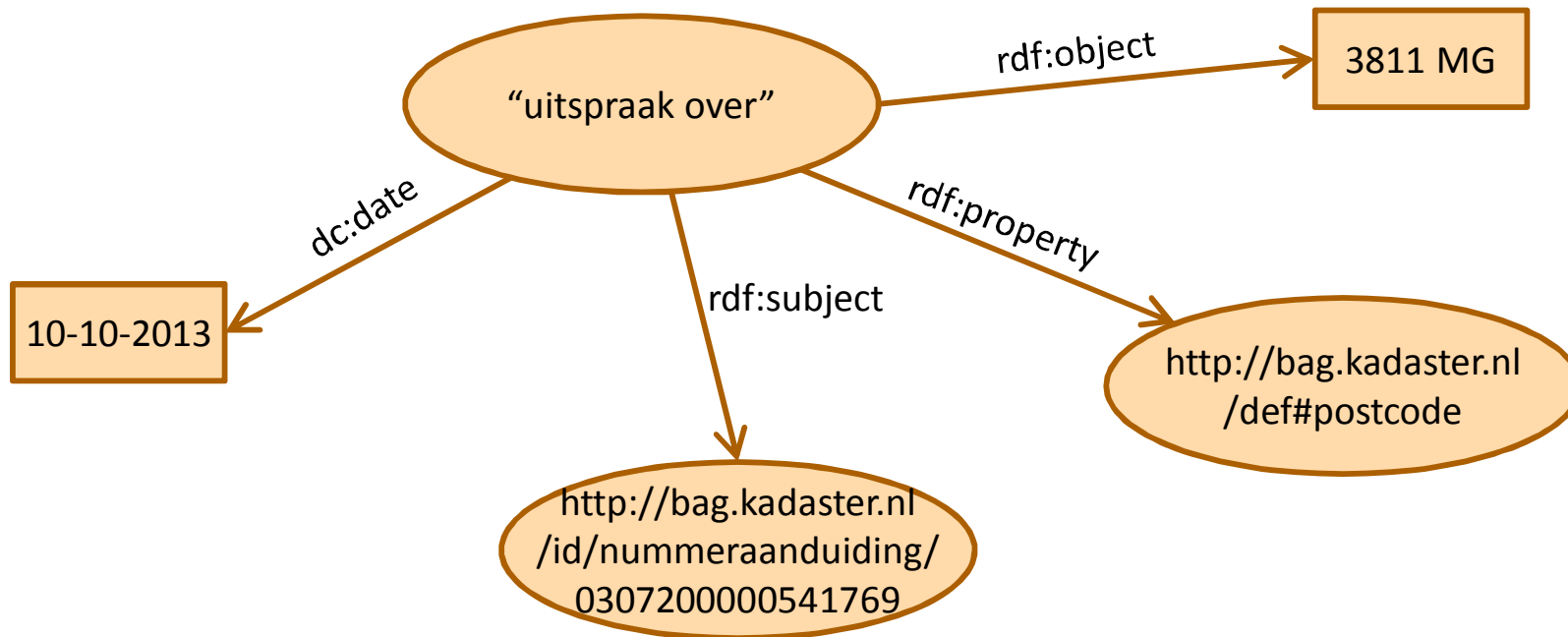
- Je moet weten welke resolver(s) er zijn om bij informatie te komen;
- Geen informatie over eigenaarschap in de URN;
- Vereist een universele afspraak over formaat en prefix.



- Enkelvoudige triples doen een **uitspraak** over een “ding”
- Zo’n uitspraak kent geen **context**:
  - Wie de uitspraak heeft gedaan?
  - Wanneer deze de uitspraak heeft gedaan?
  - Wat de geldigheid is van de uitspraak?
  - Op welke locatie de uitspraak is gedaan?
  - ...



- Oplossing #1: met reification-achtige constructies kun je dit uitbreiden
- Nadeel: complexe queries



- Oplossing #2: alle triples met dezelfde context in 1 graph
- Uitzoekpunt: performance, vereist ook een quad-store





Samen innoveren aan een duurzame digitale wereld